



EVALUASI PROGRAM BAHASA ARAB BERBASIS HOTS DAN PENEGAKAN ETIKA AKADEMIK DI PERGURUAN TINGGI

Muhammad Maulana Arif^{1*}, Sri Mulya Nengsi², Nur Atika Putri³, Erni Nurjaya⁴, Ahsan Takwim⁵, Herdah Maesarah⁶

^{1*,2,3,4,5,6} IAIN Parepare

*Email: mthmaulana@gmail.com, mulyanengsi6@gmail.com, nuratikasaid0@gmail.com,
erninurjaya8@gmail.com, ahsantakwim08@gmail.com, herdah@iainpare.ac.id

DOI: <https://doi.org/10.37081/jipdas.v6i3.5044>

Abstrak

Penelitian ini mengkaji dua permasalahan fundamental dalam evaluasi pembelajaran bahasa Arab di perguruan tinggi: (1) kesenjangan antara kurikulum Outcomes-Based Education (OBE) dengan instrumen evaluasi berbasis Lower Order Thinking Skills (LOTS), dan (2) krisis integritas akademik yang bermanifestasi dalam plagiarisme digital serta rekayasa nilai. Menggunakan pendekatan analitis-solutif dengan kajian pustaka mendalam, penelitian ini menganalisis *misalignment* OBE-LOTS melalui kerangka validity gap Biggs dan Tang (2011), efek *backwash* negatif berdasarkan teori Hughes (2003), serta mengurai krisis integritas akademik dengan menggunakan Fraud Triangle Cressey (1953) yang diintegrasikan dengan perspektif pendidikan Islam. Sebagai luaran utama, penelitian ini menghasilkan tiga prototipe instrumen evaluasi Higher Order Thinking Skills (HOTS) yang mencakup level kognitif C4-C6: instrumen Qira'ah Akademik, instrumen Kitabah Argumentatif, dan instrumen Analisis Teks Media Arab. Selain itu, disajikan pula matriks kisi-kisi evaluasi, rubrik penilaian holistik empat aspek, Draft Kode Etik Evaluator berbasis lima nilai inti (Amanah, Shidq, 'Adl, Transparansi, Akuntabilitas), instrumen deteksi plagiarisme, dan strategi redesain asesmen berbasis proyek. Temuan menunjukkan bahwa transformasi evaluasi dari paradigma LOTS menuju HOTS merupakan keniscayaan epistemologis sekaligus etis dalam konteks pendidikan bahasa Arab pascasarjana yang berorientasi pada kompetensi abad ke-21.

Kata Kunci: Evaluasi HOTS, Bahasa Arab, OBE, Etika Akademik, Kode Etik Evaluator

1. PENDAHULUAN

Kurikulum berbasis Outcomes-Based Education (OBE) menempatkan kemampuan nyata mahasiswa sebagai muara dari seluruh proses pembelajaran. Setiap komponen kurikulum—mulai dari tujuan pembelajaran, strategi instruksional, hingga instrumen evaluasi—harus selaras secara konstruktif (*constructive alignment*) untuk menjamin pencapaian kompetensi yang telah ditetapkan (Biggs & Tang, 2011). Namun, realitas empiris di banyak program Pendidikan Bahasa Arab (PBA) pascasarjana menunjukkan kesenjangan yang signifikan antara desain kurikulum OBE dan praktik evaluasi yang diterapkan.

Instrumen evaluasi yang masih mendominasi adalah instrumen berbasis Lower Order Thinking Skills (LOTS), yakni soal-soal yang hanya menguji level kognitif C1 (mengingat) dan C2 (memahami) dalam Taksonomi Bloom yang direvisi oleh Anderson dan Krathwohl (2001). Contoh konkret instrumen LOTS dalam pembelajaran bahasa Arab meliputi: hafalan kosakata secara lepas konteks, identifikasi kedudukan kata tanpa analisis makna, pilihan ganda yang menguji kaidah nahwu-sharaf secara isolatif, serta terjemahan harfiah tanpa interpretasi kontekstual. Penggunaan instrumen LOTS dalam kurikulum OBE menciptakan 'kesenjangan validitas' (*validity gap*) yang berdampak sistemik pada seluruh ekosistem pembelajaran.

Di samping persoalan epistemologis tersebut, ekosistem evaluasi pendidikan bahasa Arab juga menghadapi krisis integritas akademik yang semakin mengkhawatirkan. Dua bentuk pelanggaran yang paling dominan adalah plagiarisme digital dan rekayasa nilai. Era digital telah membuka peluang kecurangan yang sebelumnya tidak terbayangkan melalui tools kecerdasan buatan, sementara sistem



tata kelola akademik yang lemah memperparah kondisi ini. Dalam perspektif pendidikan Islam, kecurangan akademik bertentangan secara mendasar dengan nilai-nilai amanah dan kejujuran (shidq) yang menjadi fondasi etika seorang pelajar Muslim (Al-Ghazali, 2005).

Peraturan Menteri Pendidikan dan Kebudayaan Republik Indonesia Nomor 3 Tahun 2020 tentang Standar Nasional Pendidikan Tinggi (SN-Dikti) secara eksplisit mewajibkan adanya kesesuaian antara Capaian Pembelajaran Lulusan (CPL), Capaian Pembelajaran Mata Kuliah (CPMK), dan instrumen asesmen. Berdasarkan latar belakang tersebut, penelitian ini bertujuan: (1) menganalisis misalignment OBE-LOTS dan efek backwash yang ditimbulkannya; (2) mengembangkan prototipe instrumen evaluasi HOTS untuk pembelajaran bahasa Arab akademik; (3) menyusun matriks kisi-kisi dan rubrik penilaian holistik; serta (4) merumuskan kode etik evaluator dan strategi implementasi yang komprehensif.

2. METODOLOGI PENELITIAN

Penelitian ini menggunakan pendekatan analitis-solutif (analytical-solutive approach) berbasis kajian pustaka (library research) yang bersifat mendalam dan sistematis. Pendekatan ini dipilih karena tujuan penelitian adalah menghasilkan prototipe instrumen evaluasi yang dapat langsung diimplementasikan, bukan sekadar memaparkan teori. Data dikumpulkan melalui studi dokumen terhadap literatur primer dan sekunder yang relevan, meliputi buku teks pendidikan bahasa Arab, jurnal ilmiah linguistik terapan, regulasi pendidikan tinggi, dan dokumen kebijakan akademik.

Analisis data dilakukan melalui empat tahap: (1) analisis kritis terhadap misalignment OBE-LOTS menggunakan kerangka constructive alignment Biggs dan Tang (2011); (2) analisis efek backwash berdasarkan teori Hughes (2003) dan Self-Determination Theory Deci dan Ryan (2000); (3) analisis krisis integritas akademik menggunakan Fraud Triangle Cressey (1953) yang diintegrasikan dengan perspektif pendidikan Islam; serta (4) sintesis solutif berupa perancangan prototipe instrumen, rubrik, kode etik, dan strategi implementasi. Validasi prototipe instrumen dilakukan melalui expert review terhadap kesesuaian level kognitif, ketepatan linguistik Arab akademik, dan tingkat autentisitas tugas.

3. HASIL DAN PEMBAHASAN

Misalignment Kurikulum OBE dan Instrumen LOTS

Analisis menunjukkan bahwa penggunaan instrumen LOTS dalam kurikulum OBE menciptakan validity gap yang sistemik. Kesenjangan ini terjadi manakala instrumen evaluasi tidak mampu mengukur kompetensi yang sesungguhnya ingin dicapai oleh kurikulum (Biggs & Tang, 2011). Dalam konteks PBA pascasarjana, kurikulum OBE mengharapkan mahasiswa mampu menganalisis wacana akademik Arab, mengevaluasi argumen ilmiah, hingga memproduksi tulisan argumentatif yang kohesif—kompetensi yang berada pada level C4-C6 dalam Taksonomi Bloom sementara instrumen LOTS sama sekali tidak mampu mengungkap ketercapaian level tersebut.

Dampak langsung dari misalignment ini adalah reduksi bahasa Arab menjadi sekadar objek hafalan rumus gramatika. Al-Khuli (2010) menegaskan bahwa pembelajaran bahasa yang otentik harus menempatkan kompetensi komunikatif sebagai pusat evaluasi, bukan sekadar akurasi gramatikal yang diuji secara terpisah dari konteks penggunaan bahasa. Lebih jauh, misalignment ini juga mencederai prinsip transparansi dan keadilan dalam penilaian: mahasiswa yang memiliki kemampuan analisis tinggi namun lemah dalam hafalan mekanik akan dirugikan, sementara mahasiswa yang piawai menghafal tanpa benar-benar memahami dapat memperoleh nilai tinggi yang menyesatkan.

Efek Backwash Negatif dari Dominasi LOTS

Dominasi instrumen LOTS menghasilkan empat efek backwash negatif yang berdampak sistemik. Pertama, deformasi motivasi belajar: mahasiswa terjebak dalam siklus menghafal-mengerjakan soal-melupakan, yang mematikan motivasi intrinsik berbasis rasa ingin tahu (Deci & Ryan, 2000). Kedua, penyempitan praktik pengajaran melalui fenomena 'teaching to the test' (Cheng, 2004; Watanabe, 2004). Ketiga, kegagalan dalam menyiapkan kompetensi 4C abad ke-21 (Critical



Thinking, Creativity, Communication, Collaboration) yang sama sekali tidak dapat diukur oleh instrumen LOTS. Keempat, erosi kepercayaan diri komunikatif akibat meningkatnya language anxiety (Krashen, 1982).

Secara keseluruhan, efek backwash negatif menciptakan lingkaran setan dalam pembelajaran: evaluasi LOTS mendorong pembelajaran hafalan → menghasilkan kompetensi semu → nilai tinggi memvalidasi sistem LOTS yang bermasalah → dan siklus berulang tanpa koreksi. Untuk memutus lingkaran ini, diperlukan transformasi paradigmatik dari evaluasi LOTS menuju evaluasi HOTS yang otentik dan berbasis kompetensi.

Analisis Krisis Integritas melalui Fraud Triangle

Fraud Triangle Cressey (1953) mengidentifikasi tiga elemen yang harus hadir secara bersamaan agar kecurangan terjadi. Dalam konteks kecurangan akademik PBA: (1) Tekanan bersumber dari grade pressure, tekanan waktu, dan tekanan kompetensi yang dialami mahasiswa pascasarjana dengan beban multiperan; (2) Kesempatan terbuka lebar melalui tools AI generatif yang membuat plagiarisme digital mudah dilakukan namun sulit dideteksi; (3) Rasionalisasi dibangun melalui narasi seperti 'semua orang melakukannya' atau 'sistem penilaiannya tidak adil'.

Perspektif pendidikan Islam menegaskan bahwa kecurangan akademik bertentangan dengan nilai amanah dan shidq. Al-Ghazali dalam Ihya' Ulum al-Din menegaskan bahwa ilmu yang diperoleh melalui kecurangan tidak membawa berkah dan tidak menghasilkan manfaat hakiki. Hadits Nabi Muhammad SAW, 'man ghashsha fa laisa minna', menegaskan posisi Islam yang tegas terhadap segala bentuk kecurangan. Analisis Fraud Triangle menunjukkan bahwa solusi komprehensif harus menyasar ketiga elemen sekaligus: mengurangi tekanan melalui redesign evaluasi, menutup kesempatan melalui instrumen autentik, dan membangun budaya akademik yang mengikis rasionalisasi.

Prototipe Instrumen Evaluasi HOTS

Tiga prototipe instrumen evaluasi HOTS dikembangkan dengan mengacu pada level kognitif C4-C6 dalam Taksonomi Bloom yang direvisi:

- a. **Instrumen Qira'ah Akademik (C4/C5).** Menggunakan teks stimulus abstrak jurnal sosiolinguistik berbahasa Arab autentik. Soal dirancang untuk menguji: identifikasi klaim utama dan jenis argumen, evaluasi relevansi metodologi penelitian dengan referensi silang teori alternatif, dan analisis piranti kohesi leksikal-gramatikal. Teks stimulus mencakup isu kontemporer seperti fenomena diglossia (Ferguson, 1959) dan dampak era digital terhadap status bahasa Arab.
- b. **Instrumen Kitabah Argumentatif (C6).** Penugasan produksi naskah opini (maqalah ra'y) minimal 300 kata tentang isu kontemporer, contohnya: 'Pengaruh Kecerdasan Buatan terhadap Sastra Arab Kontemporer: Ancaman atau Peluang?' Instruksi mewajibkan: tesis argumentatif yang jelas, minimal dua argumen utama dengan bukti, penggunaan piranti rabth al-jumal secara tepat, penyertaan counter-argument dan responsnya, serta penggunaan kosakata akademik Arab (mufradat akademiyyah).
- c. **Instrumen Analisis Teks Media Arab (C4/C5).** Menggunakan cuplikan berita/opini dari media internasional Arab (Al Jazeera, Asharq Al-Awsat). Soal menguji: analisis nada (tone) dengan identifikasi pilihan leksikal pendukung, evaluasi keberimbangan informasi menggunakan kerangka Critical Discourse Analysis (CDA), dan analisis struktur kalimat (tarakib) termasuk jumlah f'l'iyyah vs. ismiyyah beserta implikasi stilistikanya.

Tabel 1. Matriks Kisi-Kisi Evaluasi Bahasa Arab Berbasis HOTS

Lingkup Kompetensi	Indikator Utama	Level Kognitif	Aktivitas Evaluasi
Qira'ah Akademik	Menganalisis argumen utama teks jurnal sosiolinguistik	C4	Soal uraian analisis argumen dengan teks stimulus jurnal Arab
Qira'ah Akademik	Mengevaluasi	C5	Soal uraian evaluasi metodologi



Lingkup Kompetensi	Indikator Utama	Level Kognitif	Aktivitas Evaluasi
	relevansi dan keterbatasan metodologi penelitian		dengan referensi silang
Kitabah Argumentatif	Menyusun tesis argumentatif dan mengembangkan argumen utama	C6	Makalah opini 300+ kata tentang isu kontemporer Arab
Analisis Wacana Media	Mengidentifikasi nada, register, dan bias teks media Arab	C4	Soal analisis tone dan bias dari cuplikan berita Al Jazeera
Analisis Wacana Media	Mengevaluasi keberimbangan informasi menggunakan CDA	C5	Soal evaluasi kritis dengan kerangka analisis wacana kritis
Sintaksis Fungsional	Menganalisis tarakib dan implikasi stilistika dalam teks autentik	C4	Soal analisis jumlah fi'liyyah vs. ismiyyah dan efek wacana

Rubrik Penilaian Holistik Non-Kaku

Rubrik penilaian dirancang sebagai analytical scoring rubric yang mengadopsi pendekatan Jacobs et al. (1981) dan disesuaikan dengan konteks bahasa Arab akademik pascasarjana. Rubrik mengukur empat aspek utama dengan bobot berbeda:

Tabel 2. Rubrik Penilaian Holistik Evaluasi HOTS Bahasa Arab

Aspek (Bobot)	Istimewa (90-100)	Baik (75-89)	Berkembang (60-74)	Pemula (<60)
Kedalaman Analisis (30%)	Menembus asumsi tersembunyi, mengungkap implikasi makro-sosial	Melampaui deskripsi permukaan, identifikasi hubungan sebab-akibat	Pemahaman dasar, masih terbatas pada deskripsi eksplisit	Deskriptif-faktual, tanpa upaya analisis kritis
Logika & Koherensi (25%)	Argumen logis, setiap klaim didukung bukti relevan	Umumnya logis, sebagian kecil transisi kurang mulus	Ada upaya organisasi argumen, beberapa lompatan logika	Tidak terstruktur, inkonsistensi logis signifikan
Orisinalitas Gagasan (25%)	Perspektif orisinal, sintesis inovatif lintas domain	Pemikiran mandiri signifikan, kontribusi gagasan teridentifikasi	Sebagian besar parafrase tanpa kontribusi analitik signifikan	Reproduksi langsung teks tanpa pengembangan independen
Akurasi Linguistik (20%)	Kosakata akademik Arab sangat tepat,	Kosakata umumnya tepat, beberapa	Kosakata terbatas, kesalahan	Kesalahan frekuen



Aspek (Bobot)	Istimewa (90-100)	Baik (75-89)	Berkembang (60-74)	Pemula (<60)
	piranti kohesi efektif dan variatif	kesalahan minor tidak mengganggu	gramatikal mempengaruhi kejelasan	mengganggu komunikasi, piranti kohesi absen

Kode Etik Evaluator Bahasa Arab

Dokumen Kode Etik Evaluator disusun berdasarkan lima nilai inti yang mengintegrasikan etika akademik universal dengan nilai-nilai pendidikan Islam:

- (1) **Amanah (الأمانة)**: Penilaian dilakukan semata-mata berdasarkan kualitas karya akademik, tidak berdasarkan pertimbangan personal atau kepentingan lain. Landasan normatif: QS. An-Nisa' [4]:58 tentang kewajiban menyampaikan amanah kepada yang berhak.
- (2) **Kejujuran—Shidq (الصدق)**: Melaporkan temuan kecurangan secara konsisten, memberikan umpan balik jujur dan konstruktif, mengakui keterbatasan kompetensi diri, serta tidak menyembunyikan konflik kepentingan.
- (3) **Keadilan—'Adl (العدل)**: Penerapan standar penilaian yang sama tanpa diskriminasi, pemberian kesempatan setara melalui berbagai modalitas asesmen, dan proses banding yang transparan dan dapat diakses.
- (4) **Transparansi (الشفافية)**: Penyampaian rubrik penilaian sebelum penugasan dimulai, pemberian umpan balik terperinci yang menjelaskan alasan di balik skor, dan dokumentasi yang memungkinkan audit pihak ketiga.
- (5) **Akuntabilitas (المساءلة)**: Kesiapan mempertanggungjawabkan keputusan penilaian berdasarkan bukti terdokumentasi, partisipasi aktif dalam sesi kalibrasi penilaian antar evaluator, dan penerimaan hasil audit eksternal sebagai mekanisme peningkatan kualitas.

Strategi Redesain Asesmen Berbasis Proyek

Lima strategi redesign asesmen direkomendasikan untuk meminimalisir kecurangan akademik sekaligus meningkatkan kualitas evaluasi: (1) Penerapan Project-Based Assessment dengan konteks unik yang inheren sulit dipalsukan; (2) Desain penugasan yang dipersonalisasi berdasarkan pengalaman atau konteks personal mahasiswa; (3) Evaluasi proses—bukan hanya produk—melalui peer review, writing conference, dan portofolio bertahap; (4) Penggunaan teknologi deteksi yang etis dengan kebijakan yang transparan dan prosedur investigasi yang adil; serta (5) Pendidikan integritas akademik sebagai komponen kurikulum, bukan sekadar program orientasi insidental.

4. SIMPULAN

Penelitian ini menghasilkan beberapa temuan dan luaran penting. Pertama, penggunaan instrumen LOTS dalam kurikulum OBE menciptakan kesenjangan validitas yang sistemik, mereduksi bahasa Arab menjadi objek hafalan gramatikal, dan menghasilkan efek backwash negatif yang menghambat pengembangan kompetensi abad ke-21. Kedua, analisis Fraud Triangle menunjukkan bahwa krisis integritas akademik memerlukan solusi yang menyasar tiga elemen sekaligus: tekanan, kesempatan, dan rasionalisasi—yang dalam perspektif Islam merupakan pengkhianatan terhadap amanah keilmuan.

Sebagai luaran solutif, penelitian ini menyajikan tiga prototipe instrumen HOTS (Qira'ah C4/C5, Kitabah C6, dan Analisis Media C4/C5), matriks kisi-kisi evaluasi tujuh kompetensi, rubrik penilaian holistik empat aspek, Draft Kode Etik Evaluator berbasis lima nilai inti, instrumen deteksi plagiarisme enam indikator, dan lima strategi redesign asesmen berbasis proyek. Transformasi evaluasi dari paradigma LOTS menuju HOTS merupakan keniscayaan epistemologis sekaligus etis dalam konteks pendidikan bahasa Arab pascasarjana yang berorientasi pada kompetensi abad ke-21.

Rekomendasi utama: (1) Program studi PBA segera melakukan assessment audit dan menetapkan target rasio HOTS minimal 70%; (2) Dosen menginternalisasi kode etik evaluator dan mengembangkan kompetensi merancang instrumen autentik; (3) Pimpinan institusi memberlakukan kebijakan integritas akademik komprehensif disertai mekanisme audit independen; (4) Peneliti



selanjutnya melakukan uji validitas dan reliabilitas empiris terhadap prototipe instrumen yang dikembangkan.

5. DAFTAR PUSTAKA

- Al-Ghazali, A. H. M. (2005). *Ihya' Ulum al-Din* (Jilid 1). Dar al-Minhaj.
- Al-Khuli, M. A. (2010). *Asalib Tadris al-Lughah al-Arabiyyah* (Edisi ke-3). Dar al-Falah.
- Anderson, L. W., & Krathwohl, D. R. (Eds.). (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's educational objectives*. Addison Wesley Longman.
- Biggs, J., & Tang, C. (2011). *Teaching for quality learning at university* (4th ed.). Open University Press/McGraw-Hill.
- Cheng, L. (2004). The washback effect of a public examination change on teachers' perceptions toward their classroom teaching. In L. Cheng, Y. Watanabe, & A. Curtis (Eds.), *Washback in language testing* (pp. 147-170). Lawrence Erlbaum.
- Cressey, D. R. (1953). *Other people's money: A study in the social psychology of embezzlement*. Free Press.
- Deci, E. L., & Ryan, R. M. (2000). The 'what' and 'why' of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, 11(4), 227-268. <https://doi.org/10.1023/A:1026404810133>
- Ferguson, C. A. (1959). Diglossia. *Word*, 15(2), 325-340. <https://doi.org/10.1080/00437956.1959.11659702>
- Hughes, A. (2003). *Testing for language teachers* (2nd ed.). Cambridge University Press.
- Jacobs, H. L., Zinkgraf, S. A., Wormuth, D. R., Hartfiel, V. F., & Hughey, J. B. (1981). *Testing ESL composition: A practical approach*. Newbury House.
- Kemendikbud. (2020). *Peraturan Menteri Pendidikan dan Kebudayaan Nomor 3 Tahun 2020 tentang Standar Nasional Pendidikan Tinggi*. Kementerian Pendidikan dan Kebudayaan RI.
- Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Pergamon Press.
- Partnership for 21st Century Learning. (2019). *Framework for 21st century learning*. Battelle for Kids.
- Stiggins, R. J. (2001). *Student-involved classroom assessment* (3rd ed.). Merrill Prentice Hall.
- Watanabe, Y. (2004). Teacher factors mediating washback. In L. Cheng, Y. Watanabe, & A. Curtis (Eds.), *Washback in language testing* (pp. 129-146). Lawrence Erlbaum.